

Some Problems in the Mechanical Translation of German

Leonard Brandwood, Birkbeck College, London, England*

I. RELATIVE CLAUSES

The problems discussed are those of syntactical ambiguity and multimeaning in translating relative pronouns from German to English. The former, which is of concern for the English word order, arises from the coexistence in German of homomorphous inflections and variable word order, the latter from this combined with gender dissimilarities in the two languages. Some statistics are given of the frequency with which such ambiguities were encountered in scientific texts, and some possible solutions or partial solutions discussed.

OUR CONCERN will be primarily with the problems of word order and multimeaning, and with these not in all their aspects — which would be too vast a subject for a short article — but only in connection with one particular part of sentence structure, the relative clause.

Besides relative adverbs, such as worin, darin, etc., which cause no difficulty, German uses three words to introduce relative clauses — der, welcher and was. Certain grammatical forms of these are common to two cases, with the result that the syntactical function of such forms is ambiguous, the types of ambiguity being three in number.

1. The masculine singular nominative of der and welcher is identical to the feminine singular dative.
2. The masculine singular accusative of welcher is identical to the dative plural.
3. The nominative form is identical to the accusative in the feminine singular of der and welcher, in the neuter singular of der, welcher and was, and in the plural (all genders) of der and welcher.

The first two types are rare in comparison with the third, the first especially so, because it can arise only when the relative pronoun is not preceded by a preposition. If it is preceded by a preposition, it is thereby denoted as the feminine dative, since no preposition is constructed with the nominative case. On the other hand,

without a preposition the feminine dative form is very seldom encountered. For instance, in samples of text amounting to about 30,000 words it never occurred once, while the masculine nominative occurred over 50 times. In the second type of ambiguity the dative form is likewise rare without a preposition, preference being given to the formally distinct denen. The presence of a preposition, on the other hand, does not solve the problem as before, unless the preposition is of the type which can be constructed either with accusative or with dative, but not with both.

However, we need not continue to discuss the solution of these first two types, since it will be contained in that of the main problem, the distinction of nominative from accusative in the feminine and neuter singular and the plural of all genders.

In English the functions of subject and direct object in a relative clause are indicated by the fact that, when the relative pronoun is the subject, the direct object is separated from it by the verb, e.g.,

Animals which eat men.

When the relative pronoun is the direct object, the subject occurs on the same side of the verb, e.g.,

Animals which men eat.

In German this distinction cannot be made by the position of the verb because of the rule that in a subordinate clause the verb must normally come at the end. How, then, are we to determine the function of the relative pronoun in those instances where the form of the relative provides no help?

* Now at the University of Manchester, Manchester, England.

1. The first step is to look at what follows immediately after the relative pronoun, leaving particles and the like out of consideration. If what follows is not a substantive, or if it is, but its form excludes the possibility of it being nominative, then the relative pronoun may be taken to be the subject of the clause. This applies to about half of all the instances where the function of the relative pronoun is ambiguous.

2. On the other hand, if the form of the substantive following the relative pronoun can only be nominative, the relative pronoun must be the direct object. This accounts on average for another 10 per cent of the total number of instances.

3. Thirdly, the substantive following the relative might be either nominative or accusative according to its form, but is indicated as one or the other by its congruence or otherwise with the verb (for this of course the relative pronoun must be identifiably the opposite number in each case). This too applies to about 10 per cent of all instances.

The remaining 30 per cent are those where the functionally ambiguous relative pronoun is followed by an equally ambiguous substantive. It is these which pose the real problem. Consider, for example, the following two sentences

1. Wir werden die Eigenschaften solcher Felder zu untersuchen haben sowie die Bahnen, welche die Elektronen in diesen Feldern beschreiben.

(We shall have to investigate the properties of such fields, as well as the paths which the electrons describe in these fields.)

2. Aus diesem Grunde müssen die Gleichungen in einer Form vorliegen, welche die unmittelbare Verwendung dieser verallgemeinerten Koordinaten erlaubt.

(For this reason the equations must be in a form which permits the direct use of these generalized co-ordinates.)

Now it should be remarked that, in the 70 per cent cases so far solved, the relative pronoun turns out to be the subject in 56 per cent of the instances, the direct object only in 14 per cent. We would therefore expect the ratio to be reversed in the remaining 30 per cent so far unsolved; and the expectation is fulfilled, the relative pronoun being the subject in only 5 per cent, the direct object in 25 per cent. In short, if the machine interprets every functionally ambiguous relative pronoun which it has failed to

solve as the direct object of its clause, and adopts the appropriate word order, it will be wrong once in every six such instances. Judging from the frequency of relative pronouns in texts investigated, this would mean about three times in every 10,000 words. An idea of what the incorrect word order would sound like can be obtained from the verbatim translation of example 2 above.

If a more positive solution is required, it will be necessary to consider not only the relative pronoun, but also the substantive to which it refers. Identification of this substantive alone will sometimes produce a solution by enabling a relative pronoun ambiguous in respect of number as well as case to have its number determined. Provided this differs from that of the succeeding substantive, congruence with the verb will indicate which is the subject, e.g.,

3. Wir werden die Gleichungen in der Form anschreiben, welche sie bei Verwendung dieser Einheiten annehmen.

(We shall write the equations in the form which they assume when these units are used.)

If, on the other hand, the number of the relative pronoun proves to be identical with that of the following substantive, other means of arriving at a solution will be required. A dictionary for the machine must be compiled which classifies words and indicates not only which ones can be constructed together but also in what way. Thus in the sentence

4. Allerdings wird die Wirkung dieser Felder auf Elektronen, welche sie zu verschiedenen Zeiten durchlaufen, verschieden sein.

(To be sure, the effect of these fields on electrons which traverse them at different times will be different.)

"electrons" can "traverse" "fields, * but not vice versa: nor, since Wirkung may be the other word referred to by welche or sie, are "electrons" likely to "traverse" an "action." The possibility, an "action" "traverses" "electrons," is at once excluded by congruence with the verb, and so on. Similarly in sentence 3 "the equations" may "assume" a "form," but not the other way round.

How such a classification can be achieved, and, if it is achieved, whether it will provide the complete solution, are questions still to be answered.

Finally there is the question of how the various relative pronouns are to be translated.

Was may also be used to introduce a substantival clause, including direct and indirect questions, in which case its translation is always "what." It would therefore save the trouble of having to make a distinction between this and its use as a relative pronoun, if the latter too could be translated as "what." This is possible, however, only when was refers to a preceding das, though this is the most frequent type, accounting on average for about two-thirds of all instances. If was is translated as "what," the das is left untranslated. Alternatively the das can be translated by "that" and the was by "which," e.g.,

5. Auch in diesem Falle ist es notwendig bei dem anzusetzen, was als das Kernmotiv des Werkes erkannt wurde.

(In this case too it is necessary to begin with what was recognized as the central theme of the work.)

With all other types the relative was must be translated by "which," and consequently the relative distinguished from the substantival use. This is easy enough if the was is a direct interrogative — not because of the question mark at the end of the sentence, since a relative was might well occur in an otherwise interrogative sentence, but because the direct interrogative use will occur in a main clause, the relative in a subordinate clause. The problem is how to distinguish the relative from the indirect interrogative and non-interrogative use. If the clause introduced by was is the first in the sentence, was is substantival, e.g.,

6. Was Joseph zu tun hat, ist dasselbe.

(What Joseph has to do is the same.)

When the was clause is not the first in the sentence and is substantival, it can usually be recognized as such by the absence in the preceding clause of a neuter substantive to which the was could refer. This is not an infallible rule, however, because the relative may refer not to any particular word but to the preceding clause as a whole e.g.,

7. Der Wettlauf mußte unterbrochen werden, was sehr bedauert wurde.

(The race had to be interrupted, which was greatly regretted.)

A more, but not completely certain solution results from consideration of the fact that being substantival the was clause is therefore a constituent part of an adjacent clause.

Thus, for example, in the sentence

8. Für alle, die diese Ordnung vertreten, ist das entscheidend, was die Existenz dieser Gesellschaft auszeichnet.

(For all who stand for this (social) order what distinguishes the existence of this society is decisive.)

the main clause has a construction which normally requires a predicative, and this is supplied by the was clause.

On the other hand, in the sentence,

9. Unter diesen beiden Bestimmungen läßt sich alles zusammenfassen, was für die alte Generation charakteristisch ist.

(In these two definitions can be comprehended everything which is characteristic of the old generation.)

the construction of the main clause is complete without the was clause, which is thereby denoted as relative.

With welcher too it is necessary to distinguish the relative from the interrogative use, since the interrogative form is always translated by "which," the relative either by "which" or "who". The solution is similar to that for was, however, and need not be repeated.

The main problem with both welcher and der is to decide whether they are to be translated as "who" or "which." This can be done, of course, only by establishing whether the noun referred to denotes a person or a thing. As was mentioned earlier, the relative pronoun can only refer either to the last substantive occurring before it, or — if this substantive is a dependent genitive or part of a dependent prepositional phrase — to the substantive governing this. If there is more than one dependent prepositional phrase, and if these as well as the governing substantive have dependent genitives, there will be several substantives to which the relative

pronoun might refer. Such a collection is not common, however. In most instances — about 90 per cent according to our experience — there is only one substantive for the relative pronoun to refer to, with the result that there is no problem. Nor is there if there is more than one substantive, but all denote either persons or things. The problem arises only when there is a mixture of persons and things, and then only if the substantives concerned are equally capable of being referred to by the relative pronoun, having regard to gender and number.

In this latter case there are two possible solutions. If, as we previously suggested, the dictionary incorporates a system indicating which words are constructed together, reference to this will probably decide which of the substantives, when substituted for the relative pronoun, is appropriate in the context of the relative clause. Alternatively, if we are prepared to accept some loss in variety of expression plus an occasional odd-looking, but not unintelligible translation, the whole problem can be obviated by using the word "that" for all instances of der or welcher, when used as relative pronouns. Or rather all instances except those in the genitive case. When this is possessive,

that is to say; when the relative pronoun governs a noun, it can always be translated by "whose," no matter whether referring to a person or thing. When, however, as sometimes happens, the genitive case does not indicate possession, but merely arises from construction with a preposition or verb governing the genitive, it is to be translated in the same way as other instances by "that."

If the relative pronoun is preceded by a preposition, it can still be rendered by "that," the preposition then being placed immediately after the verb in the English - "This is the man with whom I went" - "This is the man that I went with." It is with those instances, however, that the occasional odd-looking translation mentioned will be likely to arise.

Where the translation "that" fails is in the non-restrictive relative clause, e.g., "Mrs. Smith told Mrs. Jones, who then went and told Mrs. Evans." In German this is common enough with *was*, but not with *der*, the similar use of which is frowned upon by some grammarians. Apart from this and one or two other exceptions, such as the case where the substantive referred to is a person's name, the translation "that" is applicable.

II. PREPOSITIONAL PHRASES

The following is a brief consideration of the difficulty in German of determining mechanically whether a prepositional substantival phrase after a substantive is dependent on it or not, the solution of which is essential for correct word order, and therefore in many cases for the meaning in the English translation. As with the relative pronoun, the conclusion to be drawn is that a complete solution to the problem is not possible solely by syntactical considerations.

THE PROBLEM discussed in the preceding section, that of identifying the word to which the relative pronoun refers, leads to the further problem of distinguishing independent and dependent prepositional phrases.

Generally speaking, if the prepositional phrase preceding the relative pronoun is independent of the substantive in front of it, then the relative will refer to the substantive in the prepositional phrase,

1. Sonst müßte die Hochfrequenzkurve oberhalb des Sprungpunktes mit der unteren Kurve übereinstimmen, welche die Abhängigkeit des gemessenen Gleichstromwiderstandes von der Temperatur angibt.

(Otherwise the high frequency curve would have to coincide above the spring point with the lower curve, which shows the dependence of the measured direct current resistance on the temperature.)

If, on the other hand, the prepositional phrase is dependent on the preceding substantive, the relative pronoun may refer either to the substantive in the prepositional phrase, as in

2. Hieraus läßt sich ferner die ursprüngliche Zusammensetzung des Urans und das heutige Verhältnis von Pb/U und Th/U in den Primärgesteinen, die als Muttergestein der Bleiminerale gelten, in guter Übereinstimmung mit den für Granite experimentell gefundenen Zahlen berechnen.

(From this, furthermore, it is possible to calculate the original composition of uranium and the present proportion of Pb/U and Th/U in the primary rocks, which are considered to represent the parent rock of the lead minerals, in close agreement with the figures found by experiment for granites.)

or to the substantive preceding the prepositional phrase, as in the following sentence:

3. Da in der Lichtoptik es rotationssymmetrische Anordnungen von brechenden Flächen sind, welche die Abbildungen vermitteln, werden wir unser Augenmerk auf rotationssymmetrische elektrische und magnetische Felder richten müssen.

(Since in optics it is the axially symmetric arrangements of refracting surfaces which mediate the images, we shall have to direct our attention to axially symmetric electric and magnetic fields.)

It might be thought that the relative pronoun here would refer to "Flächen" rather than the abstract "Anordnungen," but this is by no means certain, as may be seen from the following example:

4. Außer den γ -Strahlen ist noch eine neue Art von Teilchen vom Atomgewicht 1 vorhanden, welche die beobachteten Protonen durch elastischen Stoß auslöst.

(Besides the γ -rays there is present a new kind of particle of atomic weight 1 which releases the observed protons by elastic collision.)

It is not so much in connection with relative clauses, however, that the distinction of independent from dependent prepositional phrase is important, as in connection with word order.

In translating, for instance,

5. Wir haben darauf hingewiesen, daß die Laplacesche Gleichung für die elektronenoptischen Felder gegenüber den lichtoptischen Medien eine Einschränkung bedeuten.

the English word order varies according to whether neither, one, or both prepositional phrases are interpreted as dependent on the preceding noun, "equation": — the different versions are

1. We have referred to the fact that the Laplace equation signifies a limitation for electrooptic fields in comparison with optical media.
2. We have referred to the fact that the Laplace equation for electrooptic fields signifies a limitation in comparison with optical media.
3. We have referred to the fact that the Laplace equation for electrooptic fields in comparison with optical media signifies a limitation.

This problem arises in fact only in subordinate clauses and in the part of a main clause after the finite verb. Since in a main clause, unless it is interrogative or imperative, the finite verb must normally be the second syntactical unit, it follows that any prepositional phrase following a substantive which occurs before the verb forms a single unit with this substantive.

The most obvious method of dealing with change of word order, first proposed by Oswald and Fletcher, is to have on the English side of the program a prescribed sequence for the various syntactical units. Basically this is

- (P) S V OP -

(where each of these — excluding the verb — comprises all its dependent units — prepositional phrase, genitives, etc. *) Such a scheme

* The P in parentheses indicates that if a prepositional phrase occurs before the subject in the German, it is to be retained in the same position in the English translation.

will suffice for the majority of clauses to be translated, and, if so desired, the exceptions can be made the subject of subsidiary rules prescribing alternative syntactical patterns. The result may be a somewhat stereotyped word order in the English, but this is no great detriment in translating scientific texts, which in the German itself — as one would expect — tend to have a less varied and less complicated clause structure than in other literature. Hence in most cases the only change necessary is for the subject to be brought before the finite verb in translating a main clause with inverted order, or for the finite verb to be advanced from the end of the clause to a position immediately after the subject in a subordinate clause: that is, the sequences (P) V S O P and (P) S O P V are to be altered to that prescribed.

If this is the limit to the rearrangement of the word order, the problem of the dependent prepositional phrase will apply only to those dependent on the subject, since they are the only ones liable to be separated from their substantive by the verb. It might be thought that it would also apply to those dependent on the direct object, when this occurred at the end of the clause instead of in its more usual place immediately after the subject or the verb — that is in a sequence such as S V P O P. In accordance with the prescribed sequence the direct object has to be transferred to the position immediately after the verb. It is unnecessary, however, to determine whether the prepositional phrase following the direct object is dependent on it or not, because in either case it can be transferred along with it. Sometimes it is not desirable to follow the prescribed sequence in such instances, but this is a separate problem and does not depend on the status of the following prepositional phrase.

There are occasions, however, when it is necessary to determine whether a prepositional phrase after a direct object, is or is not dependent. These arise with verbs of perceiving and certain others such as "permitting," when the substantive which is the direct object of the main verb is also the subject of the infinitive. An illustration of this is provided by the following sentence:

6. Wir lassen eine beliebige Ebene durch die Symmetrieachse mit der x, z - Ebene des rechtwinkligen Koordinatensystems zusammenfallen.

(We let an arbitrary plane through the axis of symmetry coincide with the xz plane of the rectangular co-ordinate system.)

If we consider only the prepositional phrases which follow immediately upon a substantive, and these only in a subordinate clause or the part of a main clause after the finite verb, then they are more often dependent on the substantive than independent, the proportion being approximately 5 : 4. In the case of those that are dependent, the substantive on which they depend is — on an average —

in 20 per cent of the instances the subject,
in 25 per cent of the instances the direct object,
and in 45 per cent of the instances an independent prepositional phrase.

In the remaining 10 per cent of the instances the substantive is a predicative, an apposition, indirect object, etc. This means that, if the rearrangement of word order is restricted to the subject and to the direct object in the accusative and infinitive construction just mentioned, only about 1 in 9 of all the prepositional phrases following a substantive causes difficulty. On the basis of texts examined this is approximately 6 per 1000 words. If, however, we wish to change the order of the prepositional phrases themselves — if, for instance, in translating sentence 3 we wish to emphasize the last prepositional phrase and say

"Since in optics it is the axially symmetric arrangements of refracting surfaces which mediate the images, it is to axially symmetric electric and magnetic fields that we must direct our attention."

we shall have to examine not one or two, but all of the prepositional phrases in the sentence concerned.

Even if we adopt the easier course and have to decide whether the prepositional phrase is dependent or not only once in 9 instances, the question still remains of how this is to be done. A partial solution — investigations suggest that about half of the relevant instances may be solved — can be achieved by including in the program various makeshift rules such as the following:

If the subject is followed by a prepositional phrase but the direct object is not, make the construction passive, so that the direct object becomes the subject and the subject the agent, e.g.,

7. Im Falle b) erreicht das Potential auf der Achse im Punkte S einen Extremwert.

(In case b) an extreme value is attained by the potential on the axis at point S.)

Similarly, where necessary, turn personal constructions such as sich lassen into impersonal ones, thereby again making the subject the direct object, e.g.,

8. Wir wollen zeigen, wie sich aus dem einzelnen Lochblendenfeld die Potentialverteilung in einem aus zwei Lochblenden L 1 und L 2 zusammengesetzten System näherungsweise bestimmen läßt.

(We intend to show how from the single aperture lens field it is possible to determine the potential distribution in a system composed of two aperture lenses L 1 and L 2 by approximation.)

With certain verbs, for instance folgen, when used intransitively or in the passive voice, and providing the syntactical unit preceding the verb is directly dependent on it, the inverted German word order can be retained in the translation, e.g.,

9. Aus den Gleichungen (53) und (55) folgt durch Bildung der Rotation das Gesetz von Biot und Savart in der Form $H = \dots$

(From equation (53) and (55) follows by formation of the curl the law of Biot and Swart in the form $H = \dots$)

Likewise when a predicative adjective stands in first position, e.g.,

10. Bemerkenswert ist das Hineingreifen des Feldes durch die Blendenöffnung auf die andere Seite der Blendenelektrode.

(Noteworthy is the intrusion of the field through the diaphragm aperture to the other side of the diaphragm electrode.)

Those prepositional phrases which remain unaccounted for by this collection of rules are best regarded as independent for two reasons, a) because, on an average, of prepositional phrases following the subject only 4 are dependent on it to every 7 independent — that is, of course, excluding those instances where the subject precedes the verb in the main clause,

b) because even if the prepositional phrase is dependent on the subject and becomes separated from it in the translation, the resulting word order is in many cases quite normal — as in the translation of sentence 4 (Part I):

To be sure the effect of these fields will be different on electrons which traverse them at different times.

This scheme, besides providing only a partial solution, lacks uniformity. It would be more satisfactory to have a system of word classification on the lines suggested in the section on the relative pronoun. In this case, however, at least three factors, as well as their relative order, would have to be specified. Thus in

Allerdings wird die Wirkung dieser Felder auf Elektronen, welche sie zu verschiedenen Zeiten durchlaufen, verschieden sein.

the members of the collocation Wirkung + auf + Elektronen would be denoted as interdependent, whereas in

Wir werden die Gleichungen in der Form anschreiben, welche sie bei Verwendung dieser Einheiten annehmen.

those of the collocation Gleichung + in + Form would not. Examination of other examples suggests that even this method will not be entirely infallible, but it could be combined with the miscellany of rules previously mentioned and it has the advantage that it is applicable to all prepositional phrases, not merely those after the subject or direct object.

In conclusion it may be said that for relative clauses and prepositional phrases, as for mechanical translation in general, a comparatively few simple rules usually suffice to solve 80 per cent or 90 per cent of any particular problem. The remaining 10 per cent or 20 per cent, however, demands a much greater program for its solution. No doubt it would be possible to work out eventually a complete system — and this we should certainly endeavour to do — but it would be so complex that whether it could be used would depend on how far the design and speed of operation of electronic computers had been or could be improved. Even if a computer with sufficient storage capacity could be built, the price of perfect translation might very well be too high in terms of computer time.